International Workshop of the CarboOcean Carbon Synthesis Groups

at the Hanse Wissenschaftkolleg,
Delmenhorst, Germany
30 November – 2 December 2007

The workshop was sponsored by the Hanse Wissenschaftskolleg (HWK), and by the EU Integrated Project CarboOcean. Additional financial support was given by the International Ocean Carbon Coordination Project (IOCCP). We are very grateful to Dr. Doris Meyerdierks and collaborators of the HWK for hosting us at HWK and making the workshop successful.

Table of contents

1. Background and aim
2. Introduction
3. Data base update
4. Overview of the crossover work
5. Pacific Synthesis group PICES
6. Learning from experience with the GLODAP synthesis work
7. Crossover analysis for specific parameters
8. Some general questions and comments
9. Break-out sessions of the three groups NA, AMS and SOCS
10. Future work and time schedule
Appendix A: List of Core Cruises for North Atlantic region
Appendix B: List of participants
Group photo

1. Background and aim

One of the major objectives of the EU CARBOOCEAN integrated project is the quantification of decadal-to-centennial Atlantic and Southern Ocean carbon inventory changes. Operationally this implies the need to quantify: "the Atlantic and Southern Ocean carbon sinks, and their decadal change, through highest accuracy measurement of the changing inventories of inorganic carbon and carbon-related tracers. Atlantic and Southern Ocean data would then be integrated into a coherent global data base. " The science delivery associated with this objective is "a large scale assessment of the ocean carbon storage".

Since the previous meeting in Kiel, the first-level quality control of the data has been continuing and software for data processing has been developed. During the second half of 2007, second-level quality control has been performed on a considerable portion of the data, with the aim of producing a data product from many different data sets which is internally consistent. This is a key step towards reaching the goals of CARINA and CarboOcean. We planned to discuss the quality control work that has been performed before by different groups and group members. Experience with the

previous GLODAP synthesis project has learned that it is vitally important that the results obtained by the different methods of quality control be compared and systematically assessed. This is the only way that a consistent data product can be produced containing data from many different cruises by many different laboratories in very different regions of the world oceans. Also, all different parameters that will be quality controlled (including carbon system parameters, salinity, dissolved oxygen, major nutrients, CFCs) should be mutually consistent. This can only be done face to face with people having worked with these data.

2. Introduction

The workshop was opened by host Mario Hoppema with a short account on the history of the CARINA project and the most recent CarboOcean workshops in Iceland in 2006 and in Kiel in March 2007 that form a continuation of that. The first CARINA data synthesis meeting took place in 1999 at this same place, the Hanse Wissenschaftskolleg (HWK, Hanse Institute for Advanced Study)! Then Doris Meyerdierks of the HWK (www.h-w-k.de) welcomed on behalf of the director. HWK is a non-profit private foundation of the German states of Niedersachsen (Lower Saxony) and Bremen and the city of Delmenhorst. The main objective of the HWK is to strengthen the research potential of the universities and research institutions in the area, in particular the University of Bremen (www.uni-bremen.de) and the University of Oldenburg (www.uni-oldenburg.de), but also including e.g. the Alfred Wegener Institute, Bremerhaven. The main instrument is to provide fellowships for scientists for a stay in Delmenhorst and collaborate with nearby institutions. Fellows can be nominated or can apply their selves for a residency at HWK between 3 and 10 months.

3. Data base update

(based on talk by Bob Key)

The initial CARINA data base consisted of 30 data sets, largely in bad condition. By the time of the Iceland meeting in 2006 there were 85, and in March 2007 (Kiel meeting) 169 cruises. Since then 10 more were added and there are 20-30 pending (some late, crucial cruises still coming in).

Currently with the 2nd level data quality control we are focussing on crossover analysis, i.e. compare values in deep water for different cruises. This technique has also used in GLODAP. Doing this in density space requires good salinity data. Now we are trying to quantify errors and offsets in the CARINA data. The results from the crossover analysis are not definitive and we need firm decisions about them. To some data we may decide to add an offset. We may also decide to take out data (in the final data product).

The following tasks are still ahead:

- Incorporate new cruises and finalize cruise summary table.
- Complete initial crossover calculations and post results on the web.
- Do required additional calibration calculations (MLR, isopycnal analyses, binned averages, property-property plots, etc.). Err on safe side by not making too many corrections, not always believe computer results without additional checks.

- Finalize calibrations and create new web summary table of correction factors/amounts.
- Update readme files to include suggested corrections as part of the meta cruise data.
- Recalculate quantities (sigma x, theta, AOU, etc)

Decisions were made about the following open questions:

- Which cruises belong to more than one region?
- Final cruise ordering for each region? → Order in table will be: Ocean, Ship, Year
- Create new merged region files with calibrations applied (discard values with flag 3+4)? → Yes, 3 regional files will be created and flags 3 and 4 discarded, because users tend to ignore flags in data products.
- Fill missing vanilla values (= supporting data, oxygen, nutrients, salinity, temperature) in data products by vertical interpolation (flag 0 if appears OK in raid quality control)? → Yes, and the missing vanilla values will become part of the final data product.
- Calculate missing carbon parameters where possible (TCO_2, Alk, pH, flags as 0 or 2)? Important decision for the data product only. → Yes and the will get flag 0. Background: The Mehrbach constants appear to give consistent results in surface waters and warm waters. pH measurements have a problem of consistency, as two errors in the calculation seem to cancel out. In the deep and cold waters the measured and calculated discrete pCO_2 may differ by 500 to 700 µatm; it is not clear if this is due to a problem in the measurements or in the constants. The Roy constants may work better in cold and deep waters. Additionally, it seems that there are regional differences in the borate vs. salinity relations (Kitack Lee is working on that).
- Add calculated carbon to individual cruise files? → No data should be added to the original cruise files.
- In cases where calculated data has been reported in the individual cruise files, those will be removed.
- Distribute cruise files, metadata, data products to CDIAC, CCHDO, etc. (CLIVAR Exchange File headers).

4. Overview of the crossover work
(based on talk by Toste Tanhua)

Crossover routines, prepared by Toste, have been released to the contributors after the Kiel meeting in March 2007. Six updates have appeared so far. Some experiences are listed below:

1. Stations within the influence radius are identified.
2. If there are several crossover regions or a repeat line, areas of clustering have to be identified.
3. For the North Atlantic (NA)and Southern Ocean Carbon Synthesis (SOCS) groups, this has to be done manually (cumbersome), for the Arctic Mediterranean Seas (AMS) group a slightly different approach was necessary → Only minor changes in scripts; for cruises processed with both routines, only small differences in offsets between NA and AMS groups.

4. A different way was developed that compared all stations of cruise 2 within the influence radius of each station in cruise 1. All the „difference“ profiles were averaged and used for the weighted average and uncertainty.
   Advantage: No need for manual clustering.
   Disadvantage: No smart decisions. Regions of high variability not so easily identified. Standard deviations are smaller.
5. Differences between crossovers may be real, i.e. due to differences in oceanography (e.g. Grand Banks, 20°W in North Atlantic). Expert judgment is needed to study results from crossovers.

As regards the interpolation:

- Piecewise Cubic hermite Interpolation
- Maximum interpolation distance for depths > 1500 m; this distance was set to 1000 m
- Interpolation at every 20 m
- For density, decreasing interpolation distance was used, this turns out to be critical
- Uncertainty in salinity influences the density. Is this a problem for us? → Maybe in a few cases

Carsten Schirnick from IfM, Kiel made a data table for the crossover results. The table is online (http://carboocean.ifm-geomar.de) and results from crossover analyses (both data and figures) can be uploaded, which many participants have done already. There are 1616 crossovers in the table, 196 cruises and 12 variables; and there are 65 registered users of the website. Crossovers generally belong to a specific region, but in some cases to more than one region. This holds especially for the adjacent NA and AMS regions. There is also a download option which produces file with all offsets. It was recognized by the meeting that the website has been a very efficient and helpful tool in coordinating the efforts of several researchers in different countries, and it was suggested that we should continue to use the website also for the final decisions on recommendations for adjustments.

During the meeting it was decided that features to the interactive website must be added. In particular, a fill-in table of recommended adjustments should be implemented in which not only numbers should be entered, but rather where all evidence (figures, data, etc.) in favour of or against recommended adjustments can be uploaded. This table would then host a link to particularly relevant cross-over analyses, region-wide average distributions, etc. In a way, this table will be the final result of the 2nd level quality control.

There are different ways to determine the offsets between data/cruises within a large data base, like the one we have here. One is to determine the regional average of a certain parameter and plot it against salinity. The relationship thus obtained can be compared with the cruise in question. Another method is the multivariate crossover comparison of a single cruise against all cruises with which it has crossovers. If several investigated variables have similar offsets, this may be an indication that the offsets of that cruise are real (compared to other cruises with which it has crossover points). For this analysis it is important to have a standard deviation from the crossover analysis (minimum of three stations per crossover are needed for this).

For quantifying the offsets in a large data base, the method described by Johnson et al. (2001) may be used. Different least square models to generate set of adjustments can be used in relation to this method. The WLSQ (weighted least squares) is the preferred method; it considers the standard deviation of the offsets. Alternatively, the WDLSQ (weighted dampened least squares) also includes the weight of the variable for a particular cruise. A program was generated which calculates offsets per cruise. One can weigh the standard deviations per cruise. If cruises are far apart in time (for our data base this is > 20 years), the weight should be time dependent and this adds a larger allowable standard deviation to older cruises. Another way of adding information to the inversion is to specify weights to a specific cruise. This could reflect the difference between "master" cruise (determined as such by the group) and "regular CARINA" cruise. This information could also be extracted from the metadata of each cruise, for instance, the stated precision, CRM or not, etc.

5. Pacific Synthesis group PICES
(based on talk by Chris Sabine)

Parallel to the CarboOcean Synthesis, a Pacific Synthesis is underway with similar goals:

- Create a data base of water column CO_2 related data for the Pacific
- Produce an internally consistent data base that has gone through a second level quality control
- Bring together research groups that measure water-column CO_2 related parameters in the Pacific
- Provide a forum for working groups for data collection and analysis.
- Estimate anthropogenic CO_2 and natural variability in the Pacific from regional to basin scales

The area of this effort is the entire Pacific Ocean, including the marginal seas and the Pacific sector of the Southern Ocean. In the latter region, the data will be shared between PICES and the Southern Ocean Carbon Synthesis group. Of the regional groups, the South Pacific group is led by Chris Sabine, Bronte Tilbrook and Masao Ishii, all also involved in the Southern Ocean Carbon Synthesis group. Additionally, Bob Key and Alex Kozyr are, like in CarboOcean, involved in the data collecting, processing and data base building work of PICES. Goal is to begin compiling the database by January 2008 with a data submission deadline of January 2009. The different time schedule for CarboOcean and PICES is a potential problem, but for the South Pacific only minor problems are expected as most data from that region is already available.

6. Learning from experience with the GLODAP synthesis work
(based on talk by Chris Sabine)

Chris Sabine, and several others now involved in the CarboOcean synthesis, have gathered invaluable experience in a previous synthesis effort, GLODAP. This experience should be and will be used in the CarboOcean synthesis work. Some major lessons learned from GLODAP are:

1. CRM information and a sense of data precision are very important.
   - If the data are noisy, it is not possible to do a first level proper quality control.

- Noisy data can bias crossover results potentially leading to an incorrect assessment of offsets in other cruises if blind automated adjustments are used.
- CRMs are the best objective way to compare cruises if analyzed properly.

2. Do not get too hung up on the crossovers.
   - All cruises are not created equally – you must decide which cruises you believe more than other cruises
   - Be careful evaluating adjustments based on just a few stations for a cruise
   - The approach is not perfect. Evaluating crossovers around a box of 4 cruises rarely gave consistent results.
   - Consider complimentary approaches (isopycnal analysis, MLR, etc.)

3. Always err on the side of not making an adjustment
   - Decide on a minimum adjustment (depends on data quality)
   - Sometimes it may be better to just eliminate a cruise than force it to fit the data set.
   - If different approaches suggest a different adjustment, consider whether an adjustment is appropriate.
   - Will an adjustment improve the final scientific result?

4. Document everything
   - Documentation is tedious and boring and critically important if you want people to trust what you did
   - Cruise adjustment is an ongoing process… every time a new cruise is made in the area the adjustments will be re-evaluated and the magnitude of any adjustments will be re-examined

7. Crossover analysis for specific parameters

Ilaria Stendardo presented results and potential problems for crossover analyses done for oxygen. Her approach has been to identify master cruises and determine offsets for these. Clustering (manual identification) appeared to be sometimes tricky. As an example, the offsets for crossovers at 20-25N change with longitude. Problem area should be excluded from final crossover results. Most cruises may have excellent O_2, but not so good nutrients and DIC. Such that the master cruises might vary per parameter.

Fiz Perez showed that most crossovers for alkalinity and pH are in North Atlantic. Alkalinity offsets are mostly from -5 to +5 umol/kg. pH values first have to be brought to one scale and one temperature. Results of a minimization routine, which minimizes the sum of squares differences, were shown, which significantly improved (decreased) the offset results for alkalinity. For pH in the NA the standard deviation of the offsets is much larger than the mean (0.021 vs 0.008, respectively). By applying the same minimization procedure the pH offsets can be optimized as well.

Reiner Steinfeldt showed data from the Labrador Sea with relatively large changes in salinity over time at depth and a temporal change in CFC-12. In contrast in the tropics there is a significant decrease of salinity, but also a similar temporal change in CFC-12. With the aid of Transit Time Distributions (TTDs) an estimation of CFC increase

rates can be done. The rate of increase is dependent on various parameters, such as the age and the time of observation. The rate of increase for CFCs may be 0-20% per year and they are different for different CFCs. For quality control other methods include calculation of the saturation state of CFCs and property-property plots of different CFCs. A substantial number of cruises may need adjustment for CFCs (10%?). Use of CFC11/CFC12 ratios between cruises will be useful for quality control work.
Comment: Changes in salinity over time are only likely in high gradient regions. Elsewhere they are unlikely.

Denis Pierrot showed some results of crossovers for NA master cruises. Offsets for TCO_2 were generally within the uncertainty of the measurements. Some issues were detected, though, which could be solved through communication with data collectors.

8. Some general questions and comments

Additive or multiplicative corrections?
In GLODAP additive corrections were used, but these may give negative values for nutrients and oxygen in surface or low-oxygen waters. Multiplicative corrections do not have this problem.

Cruise numbers do not reflect any specific order at present.

It would be useful to plot offsets by year/region or other to better identify reasons for offsets and/or real change.

A noisy cruise will give a noisy crossover. Different weighting might be useful.
Much more information should be submitted to the metadata on the CRMs, (batch number, day, time, offset per day, average offset, standard deviation in offset).

There are cases, especially in the SO, where crossover points do not have the required minimum of three stations; still in regions with few cruise data any information would help.
Put crossover result in table and upload the accompanying figure, but do not enter a standard deviation.

What is a core cruise?
A core cruise will only be adjusted in the first round of adjustments, not afterwards. It had good quality data.

Can a core cruise be a core cruise for specific parameters only?
Ideally identify core cruises which are core cruises for all parameters.
Which cruises should get a final correction? All those which show up with an adjustment sensu Johnson et al. (2001) or only those where the adjustment is larger than the standard deviation (as done in GLODAP)?
The second possibility is preferred.

Please report all problems with data files to Bob Key.

9. Break-out sessions of the three groups NA, AMS and SOCS

During the break-out sessions for the regional groups cruise data were looked at and assessed, crossover analysis for possible problematic crossover points was evaluated and if possible, adjustments were decided on. Additionally, as far as possible, core cruises were identified (not for the SO as the amount of cruises is not that big that core files are necessary). In appendix A the list of core cruises (as expocodes) for the NA group is shown.

10. Future work and time schedule

Many crossover analyses have been done manually by many people contributing to the Synthesis effort. Despite of this, a significant portion of the analyses still has to be done, the more so as more cruises will be added to the data base in the next few months. Steven van Heuven (Univ. Groningen) produced software for automatic crossover analysis, including a sophisticated routine for handling clustering. It was decided to feed the cruise data to the automated software. The manually performed crossover analyses can be used for evaluating the automatically produced offsets. Regions with known high variability should receive special attention.

Another important issue is the weighting of individual cruises. Within the three regional groups cruise data will be assessed and weights given to the cruises.

Additional methods will be applied for assessing offsets. For example, for repeat sections gridded data will be created and differences determined.

The time plan is as follows:

15 February 2008:	Completion of all crossover analyses
15 April 2008:	Completion of table of recommended adjustments to cruises based on 2nd level QC
18-19 June 2008:	Final CARINA meeting in Paris, France. Agreement on cruise adjustments. Dataset will be made “final”, and the original data and the adjusted data product will be handed over to datacenters for public release. Documentation for the dataset and metadata for the justifications underlying the adjusted data product will be finalized at the meeting

--
Mario Hoppema, Bremerhaven, February 2008

APPENDIX A

List of Core Cruises for North Atlantic region

16	74DI19970807
17	74DI19980426
18	74DI20040404
26	06GA20000506
32	29HE20020304
37	18HU19930617
42	18HU19970509
59	06MT19941012
60	06MT19941115
64	06MT19970517
65	06MT19970707
66	06MT19970815
72	06MT20040311
78	35TH20020611
79	35TH20040605
81	64TR19900701
83	64TR19910408
84	33RO2003060
85	317519930704
86	33RO20050111
87	316N19970717
88	316N19970815
89	316N20030922
90	316N20031023
91	33RO19980123
92	29HE19920714
93	49MR20031106
94	323019940104

APPENDIX B

List of participants

Mario Hoppema
Alfred Wegener Institute for Polar and Marine Research
Bremerhaven, Germany
Mario.Hoppema@awi.de

Are Olsen
Geophysical Institute/Bjerknes Centre for Climate Research
University of Bergen
Bergen, Norway
Are.Olsen@gfi.uib.no

Steven van Heuven
University of Groningen
Biological Center
Haren, The Netherlands
svheuven@gmail.com

Toste Tanhua
Leibniz Institute of Marine Sciences
Marine Biogeochemistry
Kiel, Germany
ttanhua@ifm-geomar.de

Robert M. Key
A.O.S. Program
Princeton University
Princeton, U.S.A.
key@princeton.edu

Reiner Steinfeldt
Department of Oceanography,
University of Bremen
Bremen, Germany
rsteinf@physik.uni-bremen.de

Claire Lo Monaco
LOCEAN, IPSL
Universite Pierre et Marie Curie
Paris, France
lomonaco@ccr.jussieu.fr

Pete Brown
School of Environmental Sciences
University of East Anglia
Norwich, U.K.
P.J.Brown@uea.ac.uk

Fiz F. Perez
Departamento de Oceanografia
Instituto Investigaciones Marinas (CSIC)
Vigo, Spain
fiz.perez@iim.csic.es

Dorothee Bakker
School of Environmental Sciences
University of East Anglia
Norwich, U.K.
D.Bakker@uea.ac.uk

Aida F. Ríos
Departamento de Oceanografia
Instituto Investigaciones Marinas (CSIC)
Vigo, Spain
aida@iim.csic.es

Alison Macdonald
Woods Hole Oceanographic Institution
Woods Hole, MA 02543, U.S.A.
amacdonald@whoi.edu

Antón Velo
Departamento de Oceanografia
Instituto Investigaciones Marinas (CSIC)
Vigo, Spain
avelo@iim.csic.es

Douglas Wallace
Leibniz Institute of Marine Sciences
Marine Biogeochemistry
Kiel, Germany
dwallace@ifm-geomar.de

Emil Jeansson
Bjerknes Center for Climate Research
University of Bergen
Bergen, Norway
emil.jeansson@bjerknes.uib.no

Sara Jutterström
Department of Chemistry
Analytical and Marine Chemistry
Göteborg University
Göteborg, Sweden
sara.jutterstrom@chem.gu.se

Denis Pierrot
NOAA/AOML
Miami, FL 33149, U.S.A.
Denis.Pierrot@noaa.gov

Benjamin Pfeil
Bjerknes Centre for Climate Research
University of Bergen
Bergen, Norway
Benjamin.Pfeil@bjerknes.uib.no

Marta Álvarez
Instituto Mediterráneo de Estudios Avanzados - IMEDEA
CSIC - Univ. Illes Balears
Esporles, Spain
marta.alvarez@uib.es

Chris Sabine
NOAA, Pacific Marine Environmental Laboratory
Seattle, USA
Chris.Sabine@noaa.gov

Sofia Hjalmarsson
Department of Chemistry
Göteborg University
Göteborg, Sweden
sofia@chem.gu.se

Jón Ólafsson
Institute for Marine Research
Reykjavik, Iceland
jon@hafro.is

Illaria Stendardo
ETH, Institute of Biogeochemistry and Pollutant Dynamics
Zürich, Switzerland
ilaria.stendardo@env.ethz.ch

Lynne Talley
Scripps Institution of Oceanography,
La Jolla, CA 92093-0230, U.S.A.
ltalley@ucsd.edu

Arne Körtzinger
Leibniz Institute of Marine Sciences
Marine Biogeochemistry
Kiel, Germany
akoertzinger@ifm-geomar.de

Eva Falck
Geophysical Institute
University of Bergen
Bergen, Norway
Eva.Falck@gfi.uib.no

Group photo in the lobby of the Hanse Wissenschafskolleg